Available online at www.sciencedirect.com

ScienceDirect

Procedia Economics and Finance 23 (2015) 658 – 664

Procedia Economics and Finance

www.elsevier.com/locate/procedia

2nd GLOBAL CONFERENCE on BUSINESS, ECONOMICS, MANAGEMENT and TOURISM, 30-31 October 2014, Prague, Czech Republic

Flood risk as a price-setting factor in the market value of real property

Martin Cupal[a]*

[a]*Masaryk University, Faculty of Economics and Administration, Lipová 507/41a, Brno, 602 00, Czech Republic*

Abstract

Currently, flood risk can be considered as the most serious threat, mainly in areas and countries where hardly any other natural risks occur. In relation to the field of valuation and insurance, flood risk represents a significant factor entering the new valuation procedures as well as binding regulations for real property valuation. The main objective of the research was to determine whether flood risk could be considered as a price-setting factor in market price of a real estate. If so, then it would be possible to start considering it in real estate valuation methods. Statistical methods of multiple linear regression model and statistical hypothesis testing, particularly statistical signification of regression parameter representing flood risk were employed in the research. The research was performed on the housing segment of the Czech real estate market. The paper presents an estimated model with a modified set of parameters that can be used to determine the market price of a house and also determine the degree of influence factor the flood risk may have on the final market price.

Selection and/ peer-review under responsibility of Academic World Research and Education Center

Keywords: Market value; real estate; linear regression model; price-setting factor; flood risk;

1. Introduction

Various factors enter real estate market and affect the market price. The property itself and its market price or value differ according to the value of these factors . The aim of this paper is to identify the influence of these factors and to find and test new potential factors. One part of the research involved testing the factor of flood risk as a potentially significant factor of the recent period in the area, Korytarova (2010).

* Ing. et Ing. Martin Cupal, Ph.D.. Tel.: +420 605 965 377;.
E-mail address: martin.cupal@gmail.com

Selection and/ peer-review under responsibility of Academic World Research and Education Center
doi:10.1016/S2212-5671(15)00447-5

2. Market value assessment, defining a set of price-setting factors

2.1. Market value of real property framework

When necessary to estimate market value of a real property, it is strongly recommended to follow the ensuing definition (EVS 2012 "The blue book"): "The estimated amount for which the asset should exchange on the valuation date between a willing buyer and a willing seller in an arm's length transaction after proper marketing wherein the parties had each acted knowledgeably, prudently and without compulsion." Only those transactions should enter into an estimation of the market value.

It is essential that market information of real property transactions as defined above, including the market prices, enters a market value estimation of a particular real property after qualitative adjustments and statistical evaluations have been conducted.

2.2. The classical approaches to the estimation of market value

Market value is typically indicated in accordance with IVS or EVS, using three basic valuation approaches. The first one, sales comparison approach, expresses the most direct valuation approach to real estate market. It is based on comparison of other similar objects and their market prices in the same market at the same point in time. The method works best if the comparable objects are identical. But a property can never be absolutely identical to any other property, thus special factors containing differences enter this approach.

The second one is called the income approach, which is based on such property where ownership and occupation are separate. The occupation is typically based under a contract with the occupier paying the owner rent in return for the right to occupy and the owner surrenders the occupation rights for rent. For a valuation, the valuer will generally assess the property's net income, based on comparable lettings of similar properties. Assessed property's income must be capitalised or discounted to the date of valuation.

The last main approach is the cost approach, which can be defined as the cost to obtain an asset of equal utility. The price for the site will be based on the value of comparable sites, whilst the cost of the building is based on current building costs which are then depreciated to reflect age, condition and aspects of obsolescence. This approach mainly applies to properties designed for special purpose to meet specific requirements such as churches, schools and so on; they perform non-profitable function and are not usually bought and sold, Shapiro et al (2012).

2.3. An alternative way of indicating market value

If the restrictive assumption that the market value will cover only traded real estates was introduced, it would be possible to use another effective tool for estimating the market value in the framework of the sales comparison approach (market approach). Obviously, it would be possible to talk about a particular embodiment under this valuation approach, using linear regression model. A set of price-setting factors corresponds with a set of explanatory variables with the dependent variable being the market value (unrealized market price).

2.4. The set of price-setting factors

Market value of a real estate market is affected by many factors, which may act strongly or weakly. In terms of segmentation it is a residential market with family houses. When specified like this, price-setting factors for the segment can be addressed.. It holds true that properties are always unique productions and they are immovable. Thus location is supposed to be the most significant factor that can affect the market value. Property condition is another strong price-setting factor; these two affect not only the price but each may displace the other. This can be best explained on an example of a house in very poor property condition in a prime location as compared with a house in a great condition and very unprofitable and unattractive location. In theory, the price of these two houses may be the same.

Location of properties can be understood in various levels. It means location within individual settlements (center, suburbs, outskirts), location within districts and regions, location of houses in relation to generally significant cities or civic amenities. This key price-setting factor is represented by three sub-factors in this research. These are distance to the capital (variable “location 1”, notation *L1*) in kilometers beeline; number of population in the municipality where a particular house is located (variable “location II”, notation *L2*) and economical obsolescence expression – sale coefficient (defined as market price and building costs ratio). The sale coefficient decreases below 1.0 for a location with very weak marketability and grows well above 1.0, especially in locations very close to the capital or other major cities (variable “location III”, notation *L3*).

Further, price-setting factors fall within the quantitative subset, where the number of floors (*NF*) represents the number of overground floors adapted to the prevailing purpose of housing; built up area (*BA*) determines the projection of the built-up of a house in the earth plane in square meters; usable area (*UA*) means the area used for housing; floor area (*FA*) represents all square meters of floors in a house and land area (*LA*) gives square meters of built up area and adjacent land.

A very important price-setting factor of property condition (*PC*) was evaluated within individual statistically ordinal categories. The best level is represented by newly built property; slightly worse level represents buildings after reconstruction also in a very good condition; the third, worse category includes just good buildings and the last one is that of buildings before reconstruction. The accuracy and number of categories reflects the current information representation in the real estate market. The related factor to *PC* is building structure (*BS*) that expresses the material substance from which the house is built. There are only two categories, brickwork with bricks only and mixed masonry; wooden houses are rather sporadic on this market, Din et al (2001).

Availability of a garage at the family home can be quite a significant price-setting factor (*GA*), so the presence or absence of a garage is an alternative value. The influence of this factor can be substituted by a parking space next to the house.

The last price-setting factor in the research is flood risk zone (*FRZ*) to which more attention is paid in relation to the market value and which is discussed further in the next chapter.

2.5. Flood risk as price-setting factor, definition and obtaining appropriate values

Flood risk was based on historical data observed in various locations. Entries for the observed data derived from the data of geographic information systems and cartography. The values of risk consist of ordinal scale of *n*-year water, i.e. flood zone 1 to 4, where 4 is the value of the highest risk (5-year floods), 3 expresses 20-year floods, 2 means 100-year floods and 1 being an almost risk-free zone. The obligation to assess the flood risk is imposed on the EU member states by the directive of the European parliament and by the Council on assessment and management of flood risks, Cupal (2011).

It was necessary to define a point geographically determined by the appropriate point on the flood map. The exact location is obtained using JTSK coordinates used for this purpose; it describes the location using two coordinates X and Y. This location system is widely used for other purposes (sometimes obligatory, e.g. cartography) in this country and therefore other cross-sectional data relevant to this issue can be measured, Ardielli et al. (2011).

3. The market value assessment by linear regression model

3.1. Linear regression model

Linear regression model should be an *a priori* effective type of statistical model to obtain an intended estimation of continuous dependent variable, market value of real property in this case. Ordinary Least Squares (OLS) represents a common and easiest estimation method, with its properties being under the Gauss-Markov assumptions. Obviously, it has to be a multidimensional linear regression model, because the case is solved with several explanatory variables (regressors) on a single response variable.

As a matrix notation, the multidimensional linear regression model can be expressed as $\boldsymbol{y} = \boldsymbol{Xb} + \boldsymbol{e}$ and OLS estimator of regression parameters as $\boldsymbol{b} = (\boldsymbol{X}^T\boldsymbol{X})^{-1}\boldsymbol{X}^T\boldsymbol{y}$. Matrix $\boldsymbol{X}$ consists of N rows of observations and K columns referring to explanatory variables. Regression parameter estimations denote vector $\boldsymbol{b}$ as estimation of vector $\boldsymbol{\beta}$ population. Vector of dependent variable denotes $\boldsymbol{y}$ and vector $\boldsymbol{e}$ that of residuals (sample). As an all population form (not sample) model denotes $\boldsymbol{y} = \boldsymbol{X\beta} + \boldsymbol{\varepsilon}$ and $\boldsymbol{\varepsilon}$ denotes vector of random drawing from population distribution, each ε_i

being independent of other error terms. So, OLS estimator gives the best linear approximation if certain conditions are fulfilled.

The Gauss-Markov assumptions must be verified for each assembled model. The first assumption (A1) relates to the expected value of error term being zero, which means that on average, the regression line should be correct. Assumption (A3) states that all error terms have the same variance (called homoskedasticity) and assumption (A4) imposes zero correlation between different error terms, which excludes any form of autocorrelation. It is necessary to mention that the paper works purely with cross-sectional data, thus assumption A4 does not need to be tested. It implies that $E\{\varepsilon\} = 0$ and $V\{\varepsilon\} = \sigma^2 \boldsymbol{I}_N$, where $\boldsymbol{I}_N$ is N x N identity matrix. Assumption A2 implies that $\boldsymbol{X}$ and $\boldsymbol{\varepsilon}$ are independent, consequently the matrix of regressor values $\boldsymbol{X}$ does not give any information about expected values of the error terms or their variances. Under assumptions (A1) – (A4), the OLS estimator $\boldsymbol{b}$ for $\boldsymbol{\beta}$ has positive properties; in brief it is the best linear unbiased estimator.

Another factor entering the right linear regression model evaluation is the presence of multicollinearity. This means that too high correlation between two explanatory variables may lead to problems, thus technical problem with inversion of $\boldsymbol{X}^T\boldsymbol{X}$ matrix. The term multicollinearity is used to describe the problem, when an approximate linear relationship among the explanatory variables leads to unreliable regression estimates. Quantitative evaluation of this issue works with R^2_k, which denotes the squared multiple correlation coefficient between x_{ik} and the other explanatory variables. As a direct detection of multicollinearity, the variance inflation factor (VIF), where $VIF\ (b_k) = 1/(1- R^2_k)$ is often used.

Another assumption (A5) deals with normality of error terms; it means that error terms ε_i are independent drawings form normal distribution with mean equal to 0 and variance of σ^2 . It is strongly recommended to test the assumption in a numerical or graphical way. Numerical tests include for example Jarque-Bera test, Kolmogorov-Smirnov test and its adaptation, Liliefors test, Shapiro-Wilk test and others. Normality of residuals can be verified graphically, most commonly through histogram and P-P plot, Verbeek (2008).

All the applied procedures of previous assumptions and their computings will be described in chapter 3.4 on optimal model estimation.

3.2. Hypothesis testing on linear regression model

Under the Gauss-Markov assumptions and normality of error terms, OLS estimator $\boldsymbol{b}$ has normal distribution with mean and covariance matrix $\sigma^2\ (\boldsymbol{X}^T\boldsymbol{X})^{-1}$. This result can be used to test the hypothesis of unknown population regression parameters $\boldsymbol{\beta}$.

The first important statistical test is a simple t-test. Basically null hypothesis states H_0: $\beta_k = \beta^0_k$.,where it is the researcher who sets the specific value. If the null hypothesis is not true, the alternative hypothesis H_1: $\beta_k \neq \beta^0_k$ holds true. As there are no unknown values in t_k (test statistics), this can be computed from the estimate b_k and its standard error $se(b_k)$. T-value (output of regression results) represents a special form, where $t_k = b_k / se(b_k)$. This is the case when H_0: $\beta_k = 0$. If it is rejected, then b_k differs significantly from zero. This special case actually expresses statistical significance of individual parameters of b_k and also that the corresponding variable x_{ik} has a statistically significant impact on y_i. The significance measurement depends on a chosen significance level α. P-value to t-test is a common output. If p-value is smaller than α, the null hypothesis must be rejected (thus the probability supporting null hypothesis is too low) and it means in this case that b_k is a statistically significant parameter.

F-test is another important parameter in linear regression model evaluation, where test statistics F depends on R^2 statistics (goodness of fit). F-statistics is also provided as an output of regression analysis. In special cases F-test (as a model test) works with null hypothesis H_0: $\beta_1 = \beta_2 = \ldots = \beta_k = 0$. There is a possibility that individual t-tests do reject the null, while F-test (joint test) does not or vice versa. If the F-test does not reject the null hypothesis, the estimated model will be rather poor, Verbeek (2008).

All applied tests of hypothesis will be computed and assessed in chapter 3.5.

3.3. Data set and its modifications

The dimension of the data set can be expressed as N x k matrix, i.e. 150 observations and 12 basic variables. The first step of data adjustments consisted in defining statistical data type. The dependent variable *MP* belongs to real-valued (ratio scaled) data; explanatory variable *GA* (Garage availability) is a binomial categorical (nominal scale); Building structure *BS*, property condition *PC* and Flood risk zone *FRZ* are all ordinal scale; the other variables, as Location I *L1*, Location II *L2*, Location III *L3*, Number of Floors *NF*, Built up Area *BA*, Usable Area *UA*, Floor Area *FA* and Land Area *LA* are all real-valued (ratio scaled) data.

In the second step, some variables (ordinal and binomial) were assigned numerical values. Another data modification related to the logarithmic transformation of selected variables.

3.4. The estimated linear regression model for market value assessment

The process of modeling market value of a realty started with a set of all variables. Market value of real estate *MP* must be considered as a dependent one. The remaining variables should determine *MP*. The all variable model provided $adj.R^2 = 0.515$, but only 3 variables were statistically significant at 0.05 level; the assumption of normality of the residuals was slightly broken and two variables have shown strong multicollinearity. The assumption of heteroskedasticity did not prove to be true.

Following model selection (many versions were considered to reach an optimal model) one last model showing good quality and fulfilling almost all assumptions remained. The following estimated model gives market value in CZK units.

<u>MODEL 1:</u> $MP = -2\,360\,869 + 3\,L2 + 1\,450\,778\,L3 + 7\,364\,FA + 672\,131\,PC + 506\,586\,GA$

The model provided adj.$R^2 = 0.525$, with all 5 variables being statistically significant (level 0.05). The statistics detecting multicolinearity turned out well (VIF factor of each variable being very close to 1.0) but it showed substantial heteroskedasticity of residuals and normality was slightly broken again.

Based on the above mentioned findings, logarithmic transformation of certain variables was performed and importantly, the model was changed to loglinear. It means that dependent variable MP was modified to ln *(MP)*. The process of optimal model searching started again using all variables. This time the result was better, with $adj.R^2 = 0.591$ where 5 variables were statistically significant on level 0.05 (both original and log-variables); some variables had higher VIF indicating multicollineraity. It was very important, however, that the assumption of normality of residuals was confirmed and also residuals have proved as homoskedastic. Following a thorough selection process, the optimal loglinear model was defined as follows

<u>MODEL 2:</u> $\ln(MP) = 11{,}23991 + 0.29559 \ln(FA) + 0.09574 \ln(L2) + 0.22861\,PC + 0.158\,GA + 0.40813\,L3 + 0.0008\,LA - 0.00070\,L1$

This time the model provided adj.$R^2 = 0.592$; 5 explanatory variables were statistically significant (level 0.05) with other 2 being at 0.1 level. The last 2 variables were also added into the model due to the importance of their content. VIF factor of each variable was very close to 1.0, which means no multicollinearity. The assumption of normality was fulfilled and also residuals have proved as homoskedastic. All of results obviously show that model 2 should be the best.

3.5. Results and contribution of each price-setting factor to market value assessment

The resulting model 2 reached *F*-statistic of 31,874 with *p*-value equal to 0.00, thus F-test does reject the null hypothesis and the model as a whole is statistically significant with adj.$R^2 = 0.592$ (almost 60 % of *MV* variability was explained by model 2).

All simple t-tests with p-values of particular variables can be found in the following figure.

Efekt	ln Market Price Param.	ln Market Price Sm.Ch.	ln Market Price t	ln Market Price p
Abs. člen	11,23991	0,438547	25,62988	0,000000
ln Floor area	0,29559	0,080787	3,65887	0,000356
ln Location II	0,09574	0,013462	7,11218	0,000000
Property's condition	0,22861	0,032830	6,96359	0,000000
Garage available	0,15800	0,059043	2,67610	0,008324
Location III (KP)	0,40813	0,065404	6,24007	0,000000
Land area	0,00008	0,000044	1,75748	0,080992
Location I (distance to the capital)	-0,00070	0,000413	-1,70307	0,090743

Figure 1. Results of simple t-tests for all variables.

4. Flood risk as a price-setting factor

With all the tests conducted on the variable of *FRZ*, the influence on the market value seems to be weak; certainly not statistically significant. The main reason is the fact that the variable of FRZ is not completely independent of the market value explanation and therefore other price-setting factors proved to be much stronger and more significant.

Individual t-tests revealed statistical significance of flood risk as follows: p-value of the t-test of 0.72425 showed no rejection of H_0: $\beta_{FRZ} = 0$, thus the variable could not be statistically significant.

In terms of direct links of flood risk variable *FRZ* to estimate the property market value *MV* was yet tested on correlations, namely Pearson's parametric correlation (linear dependence) and Spearman nonparametric correlation. The correlations also show a tendency of dependence, i.e. whether together with the risk of flooding the market value also increases and vice versa.

Null hypothesis states that H_0: $r < 0$, thus if ↑ *FRZ*, then ↓ *MV*. The result may confirm the expected trend (negative correlation), but may not be statistically significant. The opposite trend would mean that future house owners demand sometimes even live in flood risk areas; however this does not seem statistically significant.

The result was r = -0.065 with p-value of 0.428, and correlation being very slightly negative, but not significant. Spearman's non parametric correlation showed r = - 0.075, thus the result is confirmed.

5. Conclusion

This research showed feasibility of the linear regression model under the sales comparison approach and presented evaluation of each of the price-setting factors influencing the market value of a property. Attention was paid to the factor of flood risk, whose significance on the market value was further tested.

In general, the result of the research into whether the flood risk significantly affects considerations when buying a house showed that the analyzed market does not consider it too important; what is more, it seems almost indifferent.

References

Ardielli, J., Janasova, E., (2011). Possibilities of analysing of real estate's prices in flood areas. In International Multidisciplinary Scientific *GeoConference-SGEM, vol. 3*, (pp. 1129-1136), Albena, Bulgaria.

Childs. P. D., Ott, S. H., Riddiough, T. J., (2002). Optimal valuation of noisy real assets. In *Real estate economics, vol. 30*, nr. 3, (pp. 385-414), Malden, MA.

Cupal, M., (2011). The testing of nested insurable risk presence in market price of real estate. In Podniková ekonomika a manažment, vol. 1, nr. 5, (pp.56- 60), Zilina.

Din, A., Hoesli, M., Bender, A., (2001). Environmental variables and real estate price. In *Urban studies, vol. 38*, nr. 11, (pp. 1989-2000), Mies, Switzerland.

Evans, T. A., (2012). An estimate of the accuracy of hedonic real estate valuations using the Orange County bankruptcy. *In Economica, vol. 79,* nr. 316, (pp. 703- 720), Ithaca, NY.

Korytarova, J., Hromadka, V., (2010). Assessment of the flood damages on the real estate property in the Czech Republic area, In *Agricultural economics-zemedelska ekonomika, vol. 56*, nr. 7, (pp. 317-324), Prague, Czech republic.

Narula, S. C., Wellington, J. F., Lewis, S. A., (2012). Valuating residential real estate using parametric programming. In European journal of operational research, vol. 217, nr. 1, (pp. 120-128), Richmond, Virginia.

Shapiro, E., Mackmin, D., Sams, G., (2012). Modern methods of valuation. Estates Gazette, 11 edition, Oxford.

The european group of valuer's associations – Tegova, „The european valuation standards 2012", seventh edition, http://www.tegova.org/

The international valuation standards committee – IVSC London, "The international valuation standards IVS 2011", http://www.ivsc.org/

Verbeek, M., (2008). A guide to modern econometrics, RSM Erasmus University, Wiley, 3rd edition, Rotterdam.